

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2001, 16:51:40 ; Search time 45.66 Seconds
(Without alignments)
2159.167 Million cell updates/sec

Title: US-09-830-647-1

Perfect score: 3510
Sequence: 1 MNSGAMRIIRSKGHFGGIV.....SDNLIRAFSSPSTSTFTGR 674

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTRMBL.17:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3510	100.0	674	4 Q9UBU7	Q9ubut homo sapien
2	3124	89.0	601	4 O75226	O75226 homo sapien
3	2483	70.7	676	11 Q99MU0	Q99mu0 cricetus
4	2240.5	63.8	663	11 Q9QZ41	Q9qz41 mus musculu
5	1303.5	37.1	321	11 Q9CXP2	Q9cxf2 mus musculu
6	1170	33.3	234	4 Q9YZM6	Q9yzm6 homo sapien
7	273	7.8	1695	5 Q9URR5	Q9ur55 drosophila
8	273	7.8	1695	5 Q9NKK3	Q9nk53 drosophila
9	273	7.8	1711	5 Q9VJL0	Q9vj10 drosophila
10	273	7.8	1711	5 Q9URR4	Q9ur44 drosophila
11	208	5.9	170	4 Q9H912	Q9h912 homo sapien
12	187	5.3	1790	3 Q07380	Q07380 saccharomyc
13	182	5.2	1173	2 Q9AHK6	Q9ahk6 borrelia bu
14	181.5	5.2	1578	10 Q9AV25	Q9av25 oryza sativ
15	177	5.0	1344	2 Q49545	Q49545 mycoplasma
16	175	5.0	1084	4 Q9NKG0	Q9nkg0 homo sapien
17	173	4.9	1365	2 Q49525	Q49525 mycoplasma
18	170.5	4.9	911	3 Q06704	Q06704 saccharomyc
19	170	4.8	1386	4 Q9NS87	Q9ns87 homo sapien

20	170	4.8	1516	5 Q96154	Q96154 plasmodium
21	168	4.8	1065	2 Q9AHK9	Q9ahk9 borrelia bu
22	168	4.8	1065	2 Q9AHK8	Q9ahk8 borrelia bu
23	166.5	4.7	1871	10 Q9SRD5	Q9srd5 arabidopsis
24	166	4.7	1330	6 Q97961	Q97961 vulpes vulp
25	165	4.7	849	2 Q9AHK4	Q9ahk4 borrelia bu
26	165	4.7	1119	2 Q51228	Q51228 borrelia bu
27	164.5	4.7	2062	4 Q9H231	Q9h231 homo sapien
28	164	4.7	646	5 Q96147	Q96147 plasmodium
29	164	4.7	1946	5 Q97291	Q97291 plasmodium
30	163.5	4.7	2771	5 Q26216	Q26216 plasmodium
31	163	4.6	1278	4 Q9UPP5	Q9upp5 homo sapien
32	162	4.6	1583	4 Q15045	Q15045 homo sapien
33	162	4.6	2269	5 Q26223	Q26223 plasmodium
34	161.5	4.6	1179	2 Q9AHK3	Q9ahk3 borrelia bu
35	161	4.6	957	2 Q9AHL1	Q9ahl1 borrelia bu
36	161	4.6	1300	4 Q13999	Q13999 homo sapien
37	161	4.6	1356	4 Q14707	Q14707 homo sapien
38	161	4.6	2205	5 Q9NG02	Q9ng02 dictyostell
39	161	4.6	6815	5 Q97704	Q97704 drosophila
40	158.5	4.5	1978	4 Q15154	Q15154 homo sapien
41	158.5	4.5	1979	4 Q95949	Q95949 homo sapien
42	158	4.5	891	4 Q9Y2K3	Q9y2k3 homo sapien
43	158	4.5	980	5 Q96246	Q96246 plasmodium
44	158	4.5	1271	5 Q25860	Q25860 plasmodium
45	158	4.5	3124	5 Q9GYB8	Q9gyb8 plasmodium

ALIGNMENTS

RESULT 1
ID Q9UBU7 PRELIMINARY: PRT: 674 AA.
AC Q9UBU7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ACTIVATOR OF S PHASE KINASE.
GN ASK/H37 OR DBF4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kumagai H., Sato N., Yamada M., Mahony D., Seghezzi W., Lees E.,
RA Arai K., Masai H.;
RT "A Novel Growth- and Cell Cycle-Regulated Protein, ASK, Activates
RT Human Cdc7-Related Kinase and Is Essential for G1/S transition in
RT Mammalian Cells.";
RL Mol. Cell. Biol. 0:0-0(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Jiang W., Hunter T.;
RT "Initiation of DNA Replication.";
RT "Mammalian Cdc7/Dbp4 Protein Kinase Complex is Essential for
RT Initiation of DNA Replication.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Hollingsworth R.;
RT "Use of a semi-automated yeast two-hybrid system to identify proteins
RT that interact with the human Cdc7 protein.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB028069; BAA78326.1; -;
DR EMBL; AF160249; AAD41911.1; -;
DR EMBL; AF160876; AAD4357.1; -;
DR InterPro; IPR001357; BRCT.
DR SMART; SM00292; BRCT; 1.
SQ SEQUENCE 674 AA; 76857 MW; 353FEB7E85507E5C CRC64;

Query Match 100.0%; Score 3510; DB 4; Length 674;

Best Local Similarity 100.0%; Pred. No. 4.1e-217;
Matches 674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y	1	MNSGMARHSHKGFPOGLOVKNKKNRPSLKLKTDNRPKSKCKPLNGKVFYLDLPSVY	60
y	1	MNSGMARHSHKGFPOGLOVKNKKNRPSLKLKTDNRPKSKCKPLNGKVFYLDLPSVY	60
b	1	MNSGMARHSHKGFPOGLOVKNKKNRPSLKLKTDNRPKSKCKPLNGKVFYLDLPSVY	60
y	61	SEKLOKIDKIDGVRVEEELSRDYSLYSNNKKKAKFNOTGRISPPSPSESYAETTSPI	120
y	61	SEKLOKIDKIDGVRVEEELSRDYSLYSNNKKKAKFNOTGRISPPSPSESYAETTSPI	120
b	61	SEKLOKIDKIDGVRVEEELSRDYSLYSNNKKKAKFNOTGRISPPSPSESYAETTSPI	120
y	121	PSHDSSEKSPDYCLSLSRKLLVEAKIKODHFPNSIIISNLSMGVLIHIDIRYIE	180
y	121	PSHDSSEKSPDYCLSLSRKLLVEAKIKODHFPNSIIISNLSMGVLIHIDIRYIE	180
b	121	PSHDSSEKSPDYCLSLSRKLLVEAKIKODHFPNSIIISNLSMGVLIHIDIRYIE	180
y	181	OKKKELYLLKKSSTSVADGCKRVGSGAKCTPTGKTKKPKPKYKEDMSOLYRFPYLOLTHNP	240
y	181	OKKKELYLLKKSSTSVADGCKRVGSGAKCTPTGKTKKPKPKYKEDMSOLYRFPYLOLTHNP	240
b	181	OKKKELYLLKKSSTSVADGCKRVGSGAKCTPTGKTKKPKPKYKEDMSOLYRFPYLOLTHNP	240
y	241	FINYSIQPCSPFPVOKPSSMKOKQVYKLRIDGDKYGGTSIOLOLEKKKKKCYCCCL	300
y	241	FINYSIQPCSPFPVOKPSSMKOKQVYKLRIDGDKYGGTSIOLOLEKKKKKCYCCCL	300
b	241	FINYSIQPCSPFPVOKPSSMKOKQVYKLRIDGDKYGGTSIOLOLEKKKKKCYCCCL	300
y	301	OKYEDLETHLLSECHNRNFAOSNOQYVDDYYSKYRFPYVYKOPPKKKRIKYSVGLSP	360
y	301	OKYEDLETHLLSECHNRNFAOSNOQYVDDYYSKYRFPYVYKOPPKKKRIKYSVGLSP	360
b	301	OKYEDLETHLLSECHNRNFAOSNOQYVDDYYSKYRFPYVYKOPPKKKRIKYSVGLSP	360
y	361	VSASVYLKKTBEKREVELQHSOKDCOEDDTTVKQNLKXKHOENKELLKITSPIRPHPS	420
y	361	VSASVYLKKTBEKREVELQHSOKDCOEDDTTVKQNLKXKHOENKELLKITSPIRPHPS	420
b	361	VSASVYLKKTBEKREVELQHSOKDCOEDDTTVKQNLKXKHOENKELLKITSPIRPHPS	420
y	421	NELRGINKEMSKNSMLSTAEDDIRONFTQPLFKKQECILDISHTLSNMLELRVD	480
y	421	NELRGINKEMSKNSMLSTAEDDIRONFTQPLFKKQECILDISHTLSNMLELRVD	480
b	421	NELRGINKEMSKNSMLSTAEDDIRONFTQPLFKKQECILDISHTLSNMLELRVD	480
y	481	HYKNCITQASVHNSDFSTDNSSGQPKOKSDTYLFPKAKOLEKXDLHSTFHDSCILTIINSQ	540
y	481	HYKNCITQASVHNSDFSTDNSSGQPKOKSDTYLFPKAKOLEKXDLHSTFHDSCILTIINSQ	540
b	481	HYKNCITQASVHNSDFSTDNSSGQPKOKSDTYLFPKAKOLEKXDLHSTFHDSCILTIINSQ	540
y	541	EHLTVQAKAPFHTPEEPBNECDKRNMSLPSGKIRKVKYIILGNKKKELPBNAPFKRT	6000
y	541	EHLTVQAKAPFHTPEEPBNECDKRNMSLPSGKIRKVKYIILGNKKKELPBNAPFKRT	6000
b	541	EHLTVQAKAPFHTPEEPBNECDKRNMSLPSGKIRKVKYIILGNKKKELPBNAPFKRT	6000
y	601	EFTYQENRIGCSSPVQSLDLFLVYSEKSEPLGFTSYKSGICLVLDIWEENSQNLIT	6600
y	601	EFTYQENRIGCSSPVQSLDLFLVYSEKSEPLGFTSYKSGICLVLDIWEENSQNLIT	6600
b	601	EFTYQENRIGCSSPVQSLDLFLVYSEKSEPLGFTSYKSGICLVLDIWEENSQNLIT	6600
y	661	AFSSSPSTSTFTGF 674	
y	661	AFSSSPSTSTFTGF 674	
b	661	AFSSSPSTSTFTGF 674	

RESULT	2
075226	
ID	075226
AC	075226;
DT	01-NOV-1998 (PRELIMINARY; 08; Created)
DT	01-NOV-1998 (TEMBLrel. 08; Last sequence update)
DT	01-NOV-1998 (TEMBLrel. 08; Last annotation update)
DE	MUSGC:H.RG135C18.1 PROTEIN (FRAGMENT).
OS	Homo sapiens (human); Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NCBI-TaxID=9606;	Mammalia; Eudtheria; Primates; Carnivora; Hominoidea; Homo.
NCBI-TaxID=9606;	
SEQUENCE FROM N.A.	
Rellen J., Burkhardt J.;	
"the sequence of Homo sapiens BAC clone RG135C18.";	
Submitted (JUN-1998) to the EMBL/GenBank/DDBB databases.	
[2]	
SEQUENCE FROM N.A.	
Waterston R.;	
RA Submitted (JUN-1998) to the EMBL/GenBank/DDBB databases.	

```
DR      EMBL; AC005164; AAC23786.1; -.
FT      NON_TER      1      1
SQ      SEQUENCE      601 AA; 68666 MW; 7FC90B0234C1345F CRC64;
```

Query Match	89.0%	Score 3124	DB 4	Length 601
Best Local Similarity	100.0%	Pred. No. 2e-192		
Matches 601; Conservative	0	Mismatches	0	Indels 0; Gaps 0

Qy	74	RVEEFLSKDJSYLSINKEKFAEADQIGLSPSPSEASAYAEFTSPHSHGSGSFSPT	133
Db	1	RVEEFLSDJSYLSINKEKFAEADQIGLSPSPSEASAYAEFTSPHSHGSGSFSPT	60
Qy	134	VCLSRGKLLVKAIKDHPFIPNSNLSNLSMVGCVLTHIDHRYEEDKKELYLKKS	193
Db	61	VCLSRGKLLVKAIKDHPFIPNSNLSNLSMVGCVLTHIDHRYEEDKKELYLKKS	120
Qy	194	TSVRDGRKRVGSSAOKTRTGALRKPAYKEDMSOLYRPFYLOJTNMPINTSIOKPCSP	253
Db	121	TSVRDGRKRVGSSAOKTRTGALRKPAYKEDMSOLYRPFYLOJTNMPINTSIOKPCSP	180
Qy	254	DVDRKSSMOKOTQYKRLIOTDGRKGGSTLOLQLEKKKKGYCEGLOKYEDELFTLLSE	313
Db	181	DVDRKSSMOKOTQYKRLIOTDGRKGGSTLOLQLEKKKKGYCEGLOKYEDELFTLLSE	240
Qy	314	QHRNFAOSNOTOYVDIVSKLYDFVEYEWDPKPKKRLKYSVSLSPVSAVSLKTKBOK	373
Db	241	QHRNFAOSNOTOYVDIVSKLYDFVEYEWDPKPKKRLKYSVSLSPVSAVSLKTKBOK	300
Qy	374	KVELAHSIQKDCQEDDTYKEDONFLIKETDQEKULFISPRIPHSNRLRGLENNKSNK	433
Db	301	KVELAHSIQKDCQEDDTYKEDONFLIKETDQEKULFISPRIPHSNRLRGLENNKSNK	360
Qy	434	CSMLSTAEDDIRONFTQPLKHKNOECILDISHLSLSENDLEELRDYDKHKOIAVSHVS	493
Db	361	CSMLSTAEDDIRONFTQPLKHKNOECILDISHLSLSENDLEELRDYDKHKOIAVSHVS	420
Qy	494	DFSTDNSSGSPROKSDTVLPAPKDLKEKDLHSTFYDGSGLTINSOEHLYTOQAAPHT	553
Db	421	DFSTDNSSGSPROKSDTVLPAPKDLKEKDLHSTFYDGSGLTINSOEHLYTOQAAPHT	480
Qy	554	PREEPECDPKNMDSLPDSGRHKRYKYLILGRNKELEPNAPKREFPTQOENRIGSS	613
Db	481	PREEPECDPKNMDSLPDSGRHKRYKYLILGRNKELEPNAPKREFPTQOENRIGSS	540
Qy	614	PVOSLLDLPOTSEKSEFLGFTSYTEKSGICLVLDINEPNSDMLTAPFSSPSTPYTTC	673
Db	541	PVOSLLDLPOTSEKSEFLGFTSYTEKSGICLVLDINEPNSDMLTAPFSSPSTPYTTC	600
Qy	674	F 674	
Db	601	F 601	

RESULT	3		
RT	O99MU0	PRELIMINARY;	PRT: 676 AA.
AC	O99MU0.		
DT	01-JUN-2001 (TrEMBLrel. 17,		Created)
DT	01-JUN-2001 (TrEMBLrel. 17,		Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17,		Last annotation update)
DE	DBF4/Ask		
OS	Cricetulus griseus (Chinese hamster).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;		
OC	Cricetulus.		
CC	NCBI_Taxid=10029;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21147935; PubMed=11250080;		
RA	Guo B., Lee H.;		
RT	"Cloning and characterization of Chinese hamster homologue of yeast		
RT	DBF4 (Chdbf4).";		

RL Gene 264:249-256(2001).
 DR EMBL: AF292400; AAK21856.1; -
 SO SEQUENCE 676 AA; 75849 MW; EC7EEBDD80D40B8C CRC64;

Query Match 70.7%; Score 2483; DB 11; Length 676;
 Best Local Similarity 71.6%; Pred. No. 2,8e-151;
 Matches 488; Conservative 82; Mismatches 98; Indels 14; Gaps 7;

```

QY 1 MNSGAMRHSKHGFGGQIOVKNKRNPSLSKLTDRNPEKSKCPKLMGVLYLDPSTVI 60
DB 1 MNSGAMRHSKHGFGGQIOVKNKRNPSLSKLTDRNPEKSKCPKLMGVLYLDPSTVI 60
QY 61 SEKLOKDKIDGGRVEEFLSKDISYLSNKKKAKFAQTIGRISPVSPSPSATTAETTS 120
DB 61 SEKLOKDKIDGGRVEEFLSKDISYLSNKKKAKFAQTIGRISPVSPSPSATTAETTS 120
QY 121 PSHDSSFKSPDVTCLSRGKLLVEKAIKDHPFIPNSILSNALSMGVKILHIDIRYIE 180
DB 121 PSHDSSFKSPDVTCLSRGKLLVEKAIKDHPFIPNSILSNALSMGVKILHIDIRYIE 180
QY 121 PSHDSSFKSPDVTCLSRGKLLVEKAIKDHPFIPNSILSNALSMGVKILHIDIRYIE 180
DB 121 PSHDSSFKSPDVTCLSRGKLLVEKAIKDHPFIPNSILSNALSMGVKILHIDIRYIE 180
QY 181 OKKKELYLLKSSSVSDGKRVSGAOKTGTGRLKRPVKVEDMSQLYRPYLOLTNMP 240
DB 181 OKKKELYLLKSSSVSDGKRVSGAOKTGTGRLKRPVKVEDMSQLYRPYLOLTNMP 240
QY 241 FINTSIORPCSPFDVDRSSMOKOTQVLRIOQDGDYGGTSIOLOLEKKKKGYCECCL 300
DB 241 FINTSIORPCSPFDVDRSSMOKOTQVLRIOQDGDYGGTSIOLOLEKKKKGYCECCL 300
QY 301 OKYEDLETHLLSEQRHNPFAOSNOQYVDDIVSKLVDFEVEYEDKPPKKRIKYSVGLSP 360
DB 301 OKYEDLETHLLSEQRHNPFAOSNOQYVDDIVSKLVDFEVEYEDKPPKKRIKYSVGLSP 360
QY 361 VASAVLAKTEQKEKVELQHSOKDCQEDDTYKQNF----LYKETQETEKLLFTISEP 415
DB 361 VASAVLAKTEQKEKVELQHSOKDCQEDDTYKQNF----LYKETQETEKLLFTISEP 415
QY 416 IPHPSNELRGINKKSNKSMSTAEDDIRQNTQOLPLHKKKQECILDISETL--SEMD 473
DB 416 IPHPSNELRGINKKSNKSMSTAEDDIRQNTQOLPLHKKKQECILDISETL--SEMD 473
QY 421 MYSYTGKLGKDEKAA---SMLNASEPDIKQFTQLPPCKMEQEGILVSEHKLINRND 477
DB 421 MYSYTGKLGKDEKAA---SMLNASEPDIKQFTQLPPCKMEQEGILVSEHKLINRND 477
QY 474 LELRLVDHYKCNIOASVHSPDSTNSGQPKQKSDTVLPFAKDLKEKDLHSITFHDGL 533
DB 474 LELRLVDHYKCNIOASVHSPDSTNSGQPKQKSDTVLPFAKDLKEKDLHSITFHDGL 533
QY 478 LEO-RVGCSSVGPNSCVSHLSPENSLPQPKLADTHTFSKADLOEKDOLHVFPHDSOL 536
DB 478 LEO-RVGCSSVGPNSCVSHLSPENSLPQPKLADTHTFSKADLOEKDOLHVFPHDSOL 536
QY 534 ITTINSQEHLLVOAKAPHTPPEEPNECDFKNMDSLPSGKIHRRKYLILGRNRKENTLEPN 593
DB 534 ITTINSQEHLLVOAKAPHTPPEEPNECDFKNMDSLPSGKIHRRKYLILGRNRKENTLEPN 593
QY 537 VTLNMSKEQLLYKAGTP--SCGPQGPNECDTENTDNLPGKIORKVRLLDGO-KKKNVDP 594
DB 537 VTLNMSKEQLLYKAGTP--SCGPQGPNECDTENTDNLPGKIORKVRLLDGO-KKKNVDP 594
QY 594 AEPD-KRTEPTIOEENRICSSPVOSLLDLFQTSSEKSEFLGFTSTYTERSGICNVLDIWE 652
DB 594 AEPD-KRTEPTIOEENRICSSPVOSLLDLFQTSSEKSEFLGFTSTYTERSGICNVLDIWE 652
QY 653 ENSDULLTAFFSSPSTSTGTG 674
DB 653 ENSDULLTAFFSSPSTSTGTG 674
QY 655 ENSSSLSTFFSSPSASTGTG 676
DB 655 ENSSSLSTFFSSPSASTGTG 676

```

RESULT 4

090241 PRELIMINARY; PRT; 663 AA.
 AC 090241;
 DR 01-MAY-2000 (Tremblrel. 13, Created)
 DR 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DR 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE DBP4-RELATED PROTEIN.
 GN ASK OR MUDBF4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN 111
 RP SEQUENCE FROM N.A.

RX MEDLINE=99444905; PubMed=10517317;
 RA Lepke M., Puettler V., Strab C., Kneissl M., Berger C., Hoehn K.,
 RA Nanda I., Schmid M., Grunnt F.,
 RT "Identification, characterization and chromosomal localization of the
 RT cognate human and murine DBP4 genes.";
 RL Mol. Gen. Genet. 262:220-229(1999).
 DR EMBL: AJ003132; CAB56847.1; -
 DR MGD: MGI:1351328; Ask.
 SO SEQUENCE 663 AA; 74175 MW; 72E05CB87C3B1650 CRC64;

Query Match 63.8%; Score 2240.5; DB 11; Length 663;
 Best Local Similarity 66.6%; Pred. No. 9.4e-136;
 Matches 451; Conservative 76; Mismatches 133; Indels 17; Gaps 8;

```

QY 1 MNSGAMRHSKHGFGGQIOVKNKRNPSLSKLTDRNPEKSKCPKLMGVLYLDPSTVI 60
DB 1 MNSGAMRHSKHGFGGQIOVKNKRNPSLSKLTDRNPEKSKCPKLMGVLYLDPSTVI 60
QY 61 SEKLOKDKIDGGRVEEFLSKDISYLSNKKKAKFAQTIGRISPVSPSPSATTAETTS 120
DB 61 SEKLOKDKIDGGRVEEFLSKDISYLSNKKKAKFAQTIGRISPVSPSPSATTAETTS 120
QY 121 PSHDSSFKSPDVTCLSRGKLLVEKAIKDHPFIPNSILSNALSMGVKILHIDIRYIE 180
DB 121 PSHDSSFKSPDVTCLSRGKLLVEKAIKDHPFIPNSILSNALSMGVKILHIDIRYIE 180
QY 121 PSHDSSFKSPDVTCLSRGKLLVEKAIKDHPFIPNSILSNALSMGVKILHIDIRYIE 180
DB 121 PSHDSSFKSPDVTCLSRGKLLVEKAIKDHPFIPNSILSNALSMGVKILHIDIRYIE 180
QY 181 OKKKELYLLKSSSVSDGKRVSGAOKTGTGRLKRPVKVEDMSQLYRPYLOLTNMP 240
DB 181 OKKKELYLLKSSSVSDGKRVSGAOKTGTGRLKRPVKVEDMSQLYRPYLOLTNMP 240
QY 241 FINTSIORPCSPFDVDRSSMOKOTQVLRIOQDGDYGGTSIOLOLEKKKKGYCECCL 300
DB 241 FINTSIORPCSPFDVDRSSMOKOTQVLRIOQDGDYGGTSIOLOLEKKKKGYCECCL 300
QY 301 OKYEDLETHLLSEQRHNPFAOSNOQYVDDIVSKLVDFEVEYEDKPPKKRIKYSVGLSP 360
DB 301 OKYEDLETHLLSEQRHNPFAOSNOQYVDDIVSKLVDFEVEYEDKPPKKRIKYSVGLSP 360
QY 361 VASAVLAKTEQKEKVELQHSOKDCQEDDTYKQNF----LYKETQETEKLLFTISEP 415
DB 361 VASAVLAKTEQKEKVELQHSOKDCQEDDTYKQNF----LYKETQETEKLLFTISEP 415
QY 421 MYSYTGKLGKDEKAA---SMLNASEPDIKQFTQLPPCKMEQEGILVSEHKLINRND 477
DB 421 MYSYTGKLGKDEKAA---SMLNASEPDIKQFTQLPPCKMEQEGILVSEHKLINRND 477
QY 474 LELRLVDHYKCNIOASVHSPDSTNSGQPKQKSDTVLPFAKDLKEKDLHSITFHDGL 533
DB 474 LELRLVDHYKCNIOASVHSPDSTNSGQPKQKSDTVLPFAKDLKEKDLHSITFHDGL 533
QY 478 LEO-RVGCSSVGPNSCVSHLSPENSLPQPKLADTHTFSKADLOEKDOLHVFPHDSOL 536
DB 478 LEO-RVGCSSVGPNSCVSHLSPENSLPQPKLADTHTFSKADLOEKDOLHVFPHDSOL 536
QY 534 ITTINSQEHLLVOAKAPHTPPEEPNECDFKNMDSLPSGKIHRRKYLILGRNRKENTLEPN 593
DB 534 ITTINSQEHLLVOAKAPHTPPEEPNECDFKNMDSLPSGKIHRRKYLILGRNRKENTLEPN 593
QY 537 VTLNMSKEQLLYKAGTP--SCGPQGPNECDTENTDNLPGKIORKVRLLDGO-QANNEPSELDK 596
DB 537 VTLNMSKEQLLYKAGTP--SCGPQGPNECDTENTDNLPGKIORKVRLLDGO-QANNEPSELDK 596
QY 594 KRTREPTIOEENRICSSPVOSLLDLFQTSSEKSEFLGFTSTYTERSGICNVLDIWE 657
DB 594 KRTREPTIOEENRICSSPVOSLLDLFQTSSEKSEFLGFTSTYTERSGICNVLDIWE 657
QY 653 ENSDULLTAFFSSPSTSTGTG 674
DB 653 ENSDULLTAFFSSPSTSTGTG 674
QY 647 LSTFFSSPSTSAFVG 663
DB 647 LSTFFSSPSTSAFVG 663

```

RESULT 5

090241 PRELIMINARY; PRT; 321 AA.
 AC 090241;
 DR 01-JUN-2001 (Tremblrel. 17, Created)
 DR 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DR 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE 14 DAYS EMBRYO LIVER CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
 CLONE:4432409B02, FULL INSERT SEQUENCE.

Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC LIVER;
 RC MEDLINE=21085660; PubMed=11217851;
 RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Pleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staudil P., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni T., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez S., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
 RA Wushaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontutski S.,
 RA Hayashizaki Y.,
 RA Functional annotation of a full-length mouse cDNA collection. *;
 RL Nature 409:685-690(2001).
 DR EMBL: AK014480; BAB29383.1;
 SQ SEQUENCE 321 AA: 36324 MW: 7418172CF52A4579 CRC64;

Query Match 37.18; Score 1303.5; DB 11; Length 321;
 Best Local Similarity 80.5%; Pred. No. 4.4e-76;
 Matches 248; Conservative 23; Mismatches 36; Indels 1; Gaps 1;

QY 1 MNSGAMRHSKGFPGGIGVKNENRNPRLSKLTDNRPEKSKCKPLMGKVFYLDLPSVTI 60
 DB 1 MNLETMHSHKAPLPGGIODRNEKRNPSLSKADNRLEKSKYKPLMGKVFYLDLPSVTI 60
 QY 61 SEKLDKIDLDGGRVEEFLSKDISYLSNKKKFAQTLGRISVPSPESAVTAETTSPI 120
 DB 61 CEKLDKIDLDGGRVEEFLSKDISYLSNKKKFAQTLGRISVPSPESAVTAETTSPI 120
 QY 121 PSHDSSFKSPDTCVLSRGLLVEKAIKDHDFIPNSILSNALSMGVKILHIDIRYIE 180
 DB 121 PSHDSSFKSPDTCVLSRGLLVEKAIKDHDFIPNSILSNALSMGVKILHIDIRYIE 180
 QY 181 OKKRELYLLKSSSTVBDGGRVSGAOKTGTGRLKPFVYEDMSQLYRPYLDLPSVTI 240
 DB 181 OKKRELYLLKSSSTVBDGGRVSGAOKTGTGRLKPFVYEDMSQLYRPYLDLPSVTI 240
 QY 241 FINSYLOKPSPDVCLSRGLLVEKAIKDHDFIPNSILSNALSMGVKILHIDIRYIE 300
 DB 241 FINSYLOKPSPDVCLSRGLLVEKAIKDHDFIPNSILSNALSMGVKILHIDIRYIE 300
 QY 301 OKKRELYLLKSSSTVBDGGRVSGAOKTGTGRLKPFVYEDMSQLYRPYLDLPSVTI 360
 DB 301 OKKRELYLLKSSSTVBDGGRVSGAOKTGTGRLKPFVYEDMSQLYRPYLDLPSVTI 360
 QY 300 OKKRELYLLKSSSTVBDGGRVSGAOKTGTGRLKPFVYEDMSQLYRPYLDLPSVTI 360
 DB 300 OKKRELYLLKSSSTVBDGGRVSGAOKTGTGRLKPFVYEDMSQLYRPYLDLPSVTI 360

RESULT 6
 QY2M6 PRELIMINARY: PRT; 234 AA.
 AC QY2M6;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ACTIVATOR OF S PHASE KINASE.
 GN ASK/ H37
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kumagai H., Sato N., Yamada M., Mahony D., Seghezzi W., Lees E.,
 RA Araki K., Masai H.,
 RA "A Novel Growth- and Cell Cycle-Regulated Protein, ASK, Activates
 Human Cdc7-Related Kinase and is Essential for G1/S Transition in
 RT Mammalian Cells.";
 RL Mol. Cell. Biol. 0:0-0(1999).
 DR EMBL: AB028070; BAB78327.1;
 DR InterPro: IPR001357; BRCT.
 DR SMART: SM00292; BRCT; 1.
 SQ SEQUENCE 234 AA: 26124 MW: 50357B9CF8472BPC CRC64;

Query Match 33.3%; Score 1170; DB 4; Length 234;
 Best Local Similarity 100.0%; Pred. No. 1.1e-67;
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNSGAMRHSKGFPGGIGVKNENRNPRLSKLTDNRPEKSKCKPLMGKVFYLDLPSVTI 60
 DB 1 MNSGAMRHSKGFPGGIGVKNENRNPRLSKLTDNRPEKSKCKPLMGKVFYLDLPSVTI 60
 QY 61 SEKLDKIDLDGGRVEEFLSKDISYLSNKKKFAQTLGRISVPSPESAVTAETTSPI 120
 DB 61 SEKLDKIDLDGGRVEEFLSKDISYLSNKKKFAQTLGRISVPSPESAVTAETTSPI 120
 QY 121 PSHDSSFKSPDTCVLSRGLLVEKAIKDHDFIPNSILSNALSMGVKILHIDIRYIE 180
 DB 121 PSHDSSFKSPDTCVLSRGLLVEKAIKDHDFIPNSILSNALSMGVKILHIDIRYIE 180
 QY 181 OKKRELYLLKSSSTVBDGGRVSGAOKTGTGRLKPFVYEDMSQ 227
 DB 181 OKKRELYLLKSSSTVBDGGRVSGAOKTGTGRLKPFVYEDMSQ 227

RESULT 7
 QY09R5 PRELIMINARY: PRT; 1695 AA.
 AC QY09R5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CHIEFON.
 GN CHIEF OR CHIEFON OR BG:DS09218.2 OR CG5813.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CANTON-S;
 RA Landis G.N., Tower J.;
 RA "The Drosophila chifon gene is required for chorion gene
 RT amplification and is related to the yeast Dbf4 regulator of DNA
 RT replication and cell cycle.";
 RL Development 126:0-0(1999).
 DR EMBL: AF158178; AAD48779.1;
 DR FlyBase: FB0000307; chif.
 DR InterPro: IPR000637; AT-hook.
 DR Pfam: PF02178; AT-hook; 1.
 DR SMART: SM00384; AT-hook; 1.
 SQ SEQUENCE 1695 AA: 18080 MW: 1409DD8A1587C4B0 CRC64;

Query Match 7.88; Score 273; DB 5; Length 1695;
 Best Local Similarity 21.8%; Pred. No. 3.8e-09;
 Matches 154; Conservative 100; Mismatches 212; Indels 242; Gaps 32;

QY 27 PSUKSLTDNRPEKSKCKPLMGKVFYLDLPSVTISEKLDKIDLDGGRVEEFLSKDISYLS 86
 DB 32 PKVYVIRK-----RPLCHFEFYLDICDHLAKRIESDIKALGHLGHEFLSDITFR 83

```

QY 87 ISNKEAKFAQTIGRISVPSPESAATTAETSPHSHDSSFK-----SPDYVCLSRCKL 141
DB 84 VTDPE-----VIGTSGTPTPTST-PETPTSHQQNDGSRKPNOROSRADALISRRR 137
QY 142 LVERAKIDHDPNSLSLMSWGVKLHIDIRYVIEQKKEI--YILKSSSTVBDG 199
DB 138 STGVVNSGNSPTTSLKRSYTIW-----QTDVQRIFIKRLQTELMKQYL-----E 182
QY 200 GKRKGSQAKRTG-----RLKRPYKVEDMSQLYRPFY---LQLTNMPINYSIQPCSPF 253
DB 183 GKRGGGSGTASPHIOLKQYVYKIESVKNRYPYHLIKQPDWPKIDLSSED--GAF 240
QY 254 DV-----DKPSSMOKOTYKRLIOTDGDYGGYSIQLO---LKEKKK----- 293
DB 241 RLILKSTKTKDEHSKTRKPLGSRTSQKD--KQACEAPKPLQHPSLQELKQSAIPNSPSRN 299
QY 294 -----GYCECLOKYEDELLETHLSEQRNFAOS-----NOY 324
DB 300 CREPIDSEKGGVCEICKLEYDILNHLQSKHHLFAKNNDNLALDTLLQSSADVNR 359
QY 325 QVDDIYKLVDFVE-----YEKOTP-----KKRI--KYSVGLSPVSAV-- 365
DB 360 LEEBPVSELDMDVDESLSNEELQSPRQRPSPALREKSKRITKGGHSEKFGVAVASPO 419
QY 366 -----LKTTEKEKVELQHIQKQCEDDTV-----KEQN----- 396
DB 420 TPEPKAKVQNSPQSGISELQ---QEHPTTAAPTPTNSGRRTQNSGLSPKRAMLP 476
QY 397 -FLYKETOETKRLTISEP-----IPHPS-----NELRGINEKMSKCS 435
DB 477 SSITKAVETRE-----CATPPRGKGRPNQVDSLIYKPKOKITQELQRLNGE----- 526
QY 436 MLSTAEDDIRQNF-----TQPLKHNKQECILDISEHTLSENDLELRVDRHYKNQIASVH 491
DB 527 -----AENFEPRAVPTTSSSELPTVDVROTSD-----VRKYS 563
QY 492 VSDSTDSGQPKQKSTVY-----FPADLKAKKDLHSIFTHDSG 532
DB 564 ISSASLPTSTSEATKESGSLPTSTIRKRAQAVGRRKRVGGAADVDFPQROI-----STG 617
QY 533 LITNSSEHTVQAKAPFHTPEEPNECDFKMNDSLPSCGIHKKVKI 580
DB 618 SSSSNSMQR-----FPSAPLOPE-----GPQOPKPOLKIKI 651

```

```

RESULT 8
ID 09NK53 PRELIMINARY; PRT; 1695 AA.
AC 09NK53;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CHIF PROTEIN.
GN CHIF OR BG:DS09218.2 OR CG5813.
OS Drosophila melanogaster (fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyridiidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-Y AND CN BW SP.
RA MEDLINE=9403001; PubMed=10471707;
RA Ashburner M., Miska S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazolo M., Reese M.C., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Celniker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
[2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-Y AND CN BW SP.
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoft C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Faifan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Humastli S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomoton M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon S., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sechi H., Snir E., Svitskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.U., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abbel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Boltsakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Jaisl M., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jostali D., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidenkamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
RA Zheng X.H., Zhong F.-N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003416; AAF45011.1; -
DR EMBL: AE003650; AAC22437.1; -
DR FlyBase; FBgn0000307; chif.
DR InterPro; IPR000637; AT_hook.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF02178; AT_hook_1.
DR SMART; SM00384; AT_hook_1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SQ SEQUENCE 1695 AA; 188046 MW; B3B85B35C3DAFA0 CRC64;

```

```

Query Match 7.8%; Score 273; DB 5; Length 1695;
Best Local Similarity 21.8%; Pred. No. 3,8e-09;
Matches 154; Conservative 100; Mismatches 212; Indels 242; Gaps 32;
QY 27 PSLKLTQDNRPESKCKPLMGKVFYLDLPSTYISEKIQKIDKLGAGVEFLSKDISYL 86
DB 32 PKVYIKSK-----RPLCHFKEFYLDICDHLAKRIESDKALGHLGHEFLSDIDH 83
QY 87 ISNKEAKFAQTIGRISVPSPESAATTAETSPHSHDSSFK-----SPDYVCLSRCKL 141

```

```

DB 84 VTDKPE-----VIGGTSNTPGPPST-PCTPTSHVQDNGSARKPNQROGRADALISRYRR 137
DB 142 LVEKAIKDHDFIPNSILSNLMSMGVKLIHIDIRYIEBKKEL--YLLKSSSYVADG 199
DB 138 STGVVANSNGSTPTSLKSYTIW-----QTDVAQRIRIKIOTELKQYL-----E 182
DB 200 GKRVGSAQKTRTG--RLKRPVAVEDMSOLYRPFY--LQLTNMPFINYSIQPCSPF 253
DB 183 GKRGKGGGSGTSASPHHIOQLKQYVATIESVKRNRPYHLIKOPDDMKPDLDSSE--GAF 240
DB 234 DV-----DKPSSMOKQTVKRIQTQDGYGGSSTQIQ--LKEKKKK----- 293
DB 241 RLTKRSKTKDKKSHSMTRKPLGSRTSQKD--KQAGEAKPLOHPSLQELKQSAIPNSPSRN 299
DB 294 -----GYCECCLOKYEDLETHLSEOHNFPAOS-----NOY 324
DB 300 CREPIDSEKOGGVCEICKLEVDILINILQSKDHELEAKNSDNFLADLTLIQSDAVNRF 359
DB 325 QVVDIVSKIVPEFVE-----YEKDRP--KKKRI--KTSVGSLSPVASV-- 365
DB 360 LEEPVSESLDMVDSELSNEELQSPRQSPALREKSKRTITGKHSSEKQGVAVASPO 419
DB 366 -----LKTROKKEVLEHIOHISOKDQEDDTYV-----KEON----- 396
DB 420 TPFGAKKVGNGSGSLSELQIR--QEHPTTAATPTTNSGRKKTQNSGLSPPKRAMLP 476
DB 397 -FIKKEOTETKLLFISRP-----IPHS-----NELGLNEKMSNCS 435
DB 477 SSIKVVETREE--CATPPKRGPRPNQVDSPLVKTQROTELQRLNGF----- 526
DB 436 MLSTRAEDDIRQNF--TQLPLHKKKQECILDISERHLSNDELELRVHDKYCNQOASV 491
DB 527 -----AENMFPPPTAVPTTRSSSELPYDQDQRTSD-----VGRYS 563
DB 492 VSDFTDNGSGSQPKQKSDTVL-----FPAKDKERDLHSITFDHG 532
DB 564 ISSASLPTSTSEATEKSSGLPTSLIRKRAQAVGRRKRVGAQAQDVFQRL-----STG 617
DB 533 LITNNSQOHLTVOAKAPFTPEEPNECDPKNMDSLPSKIRHKVY 580
DB 618 SSSNSNQOR-----FPSAPIQPEE-----GPOFQPKPOLKIKI 651

```

```

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunker B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodak A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Idbegyan C.,
RA Jaitai M., Kalish F., Karpen G.H., Ke Z., Kennison J.D., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Klup J., Lai Z.,
RA Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Miliushina N.V., Mobaraj C., Morris J., Moshell A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkac R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou W., Zhou X., Zhang G., Zhao Q.,
RA Gblbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RA Science 287:2185-2195(2000).
RA EMBL: AEO03650; AAF53530.1;
RA FlyBase: FBgn0000307; ch1f.
RA Interpro: IPR000637; AT-hook.
RA Pfam: PF02178; AT-hook_1.
RA SMART: SM00384; AT-hook_1.
RA SEQUENCE 1711 AA; 189216 MW; 4CF179D974817BDA CRC64;

```

Query Match 7.8%, Score 273; DB 5; Length 1711;
 Best Local Similarity 21.8%, Pred. No. 3, 8e-09;
 Matches 154; Conservative 100; Mismatches 212; Indels 242; Gaps 32;

```

DB 27 PSLKSLKTDNRPEKSKCPKMGKVYLDLPSTYISEKLOKIDKLGRAVEELKSDISYL 86
DB 32 PKYAVIKSK-----RELCHFKEYYLDICDHOALAKRIESDIKALOGHLEFFLSDDTTHF 83
DB 87 ISNKEAKFAQTIGRISPPSPESAVYTAETTSPPHSDGSSFK-----SPPTVCLSNKLT 141
DB 84 VTDKPE-----VIGGTSNTPGPPST-PCTPTSHVQDNGSARKPNQROGRADALISRYRR 137
DB 142 LVEKAIKDHDFIPNSILSNLMSMGVKLIHIDIRYIEBKKEL--YLLKSSSYVADG 199
DB 138 STGVVANSNGSTPTSLKSYTIW-----QTDVAQRIRIKIOTELKQYL-----E 182
DB 200 GKRVGSAQKTRTG--RLKRPVAVEDMSOLYRPFY--LQLTNMPFINYSIQPCSPF 253
DB 183 GKRGKGGGSGTSASPHHIOQLKQYVATIESVKRNRPYHLIKOPDDMKPDLDSSE--GAF 240
DB 234 DV-----DKPSSMOKQTVKRIQTQDGYGGSSTQIQ--LKEKKKK----- 293
DB 241 RLTKRSKTKDKKSHSMTRKPLGSRTSQKD--KQAGEAKPLOHPSLQELKQSAIPNSPSRN 299
DB 294 -----GYCECCLOKYEDLETHLSEOHNFPAOS-----NOY 324
DB 300 CREPIDSEKOGGVCEICKLEVDILINILQSKDHELEAKNSDNFLADLTLIQSDAVNRF 359
DB 325 QVVDIVSKIVPEFVE-----YEKDRP--KKKRI--KTSVGSLSPVASV-- 365
DB 360 LEEPVSESLDMVDSELSNEELQSPRQSPALREKSKRTITGKHSSEKQGVAVASPO 419
DB 366 -----LKTROKKEVLEHIOHISOKDQEDDTYV-----KEON----- 396
DB 420 TPFGAKKVGNGSGSLSELQIR--QEHPTTAATPTTNSGRKKTQNSGLSPPKRAMLP 476
DB 397 -FIKKEOTETKLLFISRP-----IPHS-----NELGLNEKMSNCS 435

```

```

Db 477 SSIVKVTREE-----CATPRGRGRPNQVDSPLIVKFORIGTELQRLNGE----- 526
QY 436 M1STAEDDR0NF-----TOLPLKKNOCICLIDISEHTISENDLELRVDHYCNIOASVH 491
Db 527 -----AENFMPEPTAVPTTRSSSELPTVDROTSD-----VGRYS 563
QY 492 VSDFTDNGSGPOKOKSDVVL-----FPAKDLKEDLHSIFTHDSG 532
Db 564 ISSASLDISTSEAEIKESSGLPTSIKRAQAVGRRKKVGGAAAOVFOROL-----STG 617
QY 533 LTTINSOEHITVQAKAPHTPPEEPNECDFKNMDSLPSGKJHRKYKI 580
Db 618 SSSSNSNOOR-----PPSAPIQPEE-----GPOPPKPOLKIKI 651

RESULT 10
QY09R4 PRELIMINARY; PRT; 1711 AA.
AC 0909R4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CHIFON-2.
GN CHIF OR CHIFON OR BG:DS09218.2 OR CG5813.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S.
RA Landis G.N., Tower J.;
RT "The Drosophila chifon gene is required for chorion gene
RT amplification and is related to the yeast Dbf4 regulator of DNA
RT replication and cell cycle.";
RL Development 126:0-0(1999).
DR EMBL; AF158179; A048780.1; -.
DR FlyBase; FBgn000307; chif.
DR InterPro; IPR000637; AT_hook.
DR Pfam; PF02178; AT_hook; 1.
DR SMART; SM00384; AT_hook; 1.
SQ SEQUENCE 1711 AA; 189250 MW; B7D0F95517C4B9A9 CRC64;

Query Match 7.8%; Score 273; DB 5; Length 1711;
Best Local Similarity 21.8%; Pred. No. 3.8e-09;
Matches 154; Conservative 100; Mismatches 212; Indels 242; Gaps 32;

QY 27 PSLKSLKTDNRPKSKCKPLMGKVFYLDLPSTVITSEKLOKIDLGGRVEFLSDISYL 86
Db 32 PRVKVTKSK-----RFLCHFKFYLDICDHQLAKRIESDIKALGHLFEFLSDITHF 83
QY 87 ISNKEAKFAQTIAGRISPPSPESAAYTAETTSPPHSHDSSFK-----SPDVCLSRGLT 141
Db 84 YVDKNE-----VIGGTSQTPPT-PTGPTSHYQOONSARKPNROGRRAAILSRVAR 137
QY 142 LVEKAIKDHDFIPNSIISNALSMGVKILIHDDIRYIEQKKEL--YLLKSSSTVRDG 199
Db 138 STVGAVNNSGNSPTTSLKSYTIW-----QTDYAOREFIKRIOTELKOYL-----E 182
QY 200 GKRVSQAQKTTG---RLKKPFVVEDMSQLYRPY---LQITNMPFINYSIOKPCSPF 253
Db 183 GKKEGGGGSSTASPHHIDIKOYVKIESYKRVNRPYHLIKOPDDWPKLIDLSSD--GAF 240
QY 254 DV-----DKPSSQOKOTQVLRIOITDGDYKGTSTIQ--LKEKKK----- 293
Db 241 RLUTSKTKTKDKHSMTRKPLKSGRTSQKD--KQAAAGEAKPQLQHSISIQELKQASALPNSPRN 299
QY 294 -----GYECCLQKYEDLETHLLSEQHRNPAQS-----NOY 324
Db 300 CREPIDSEKQGVCEICKLEYDILNIHOSKDHELFAFNSDNFLADLTIGSADVNRH 359

```

```

QY 325 QVVDIVSKLVDFVE-----YEKDTP---KKRI---KYSVSGSLSPVSAV--- 365
Db 360 LEEPESELDVDVDESLSNEILQSPRRQSPALREKSKRTIKGKSSKFOGVAVASPO 419
QY 366 -----LKTEQKEKVELQHIISOKDCOEDTV-----KEON----- 396
Db 420 TFPPAKKVQAGNSPGLSELQ--QEHPTAAAPPTTNSGRRKTONSGLSPPKRAMLP 476
QY 397 -FLYKETOETEKLLFISPP-----IPHS-----NELGLNEKMSNKS 435
Db 477 SSIVKVTREE-----CATPRGRGRPNQVDSPLIVKFORIGTELQRLNGE----- 526
QY 436 M1STAEDDR0NF-----TOLPLKKNOCICLIDISEHTISENDLELRVDHYCNIOASVH 491
Db 527 -----AENFMPEPTAVPTTRSSSELPTVDROTSD-----VGRYS 563
QY 492 VSDFTDNGSGPOKOKSDVVL-----FPAKDLKEDLHSIFTHDSG 532
Db 564 ISSASLDISTSEAEIKESSGLPTSIKRAQAVGRRKKVGGAAAOVFOROL-----STG 617
QY 533 LTTINSOEHITVQAKAPHTPPEEPNECDFKNMDSLPSGKJHRKYKI 580
Db 618 SSSSNSNOOR-----PPSAPIQPEE-----GPOPPKPOLKIKI 651

RESULT 11
QY0912 PRELIMINARY; PRT; 170 AA.
AC 090912;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CDNA FLJ13087 F15, CLONE NT2RP302099.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Magatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takehashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK023149; BAB14431.1; -.
SQ SEQUENCE 170 AA; 18392 MW; C93DEF02A59B282E CRC64;

Query Match 5.9%; Score 208; DB 4; Length 170;
Best Local Similarity 36.4%; Pred. No. 3.2e-06;
Matches 55; Conservative 22; Mismatches 44; Indels 30; Gaps 4;

QY 21 KNEKNRPSLSLKTDRNRPKSKCKPLMGKVFYLDLPSTVITSEKLOKIDLGGRVEFLS 80
Db 35 KCKNSPFGARK-----HPSGKSFYLDLPAGNKLQFLTGAIOQLGGVIGFGLS 82
QY 81 KDISYLSNKKKFAKFOIQLGRI-----SPVSPESAAYTAETTSPPHSHDSSFKSPDTC 135
Db 83 KEYSYVSSRRREK-RESSGKSHRGCPSPEVRETVETAMWDPKSHRPSKRPVDSVP 141
QY 136 LSRGKLVEKAIKDHDFIPNSIISNALSMG 166
Db 142 LSRGKELLQKAIK-----NOVSMG 160

RESULT 12
QY07380 PRELIMINARY; PRT; 1790 AA.
AC 007380;
ID 007380;
DB 007380; P89892;

```

DT 01-NOV-1996 (TReMBrel. 01, Created)
 DT 01-NOV-1996 (TReMBrel. 01, Last sequence update)
 DT 01-JUN-2001 (TReMBrel. 17, Last annotation update)
 DE HPROTHERICAL 206.5 KDA PROTEIN YD1058W.
 GN USOI OR YD1058W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Blocker H., Brandt P.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 274105; CA98621.1; -;
 DR EMBL; 274105; CA98620.1; -;
 DR SGD; 50002216; USOI.
 DR InterPro; IPR002017; Spectrin.
 SQ SEQUENCE 1790 AA; 206450 MW; 90062544F55A52EE CRC64;

Query Match 5.3%; Score 187; DB 3; Length 1790;
 Best Local Similarity 19.9%; Pred. No. 0.0013;
 Matches 150; Conservative 133; Mismatches 273; Indels 198; Gaps 31;

QY 2 NSGAMRIHSGHFGGIGVKNKRNPSLKLKTDRPEKSKCKPLMKVFPYLDLPSTVTS 61
 DB 978 NMSLLKRAVEESKNSSTQLSNLIKIDMSQKEN-----PQIRGSTIEKN 1023
 QY 62 -EKIQOKIDKIGVNEFLSK-----DISYLINKEAKFA--QTIGRISPVSP 108
 DB 1024 ISQLKRTISDLEQTKREIKSDSSKDEYESQISLKKELTATANDENVNKISLTYT 1083
 QY 109 ESAVYAEFTTSPHSDSSKSPDVCLSRGLVKAIKDHFIPNSILSNALSMGVK 168
 DB 1084 REELAEELAA-----YKNLNELETRELS-----EKLKE--VKENE----- 1119
 QY 169 IAHIDIRYIEOK---KKELYLLKSSSTSVRDGKRVGSAQKTRGRKLPYKVED 224
 DB 1120 -EHLKEEIQLEKATEYKQOLNSLANLESLEKEHDLAQLKYEQIANKEROYNEE 1178
 QY 225 MSQLYRPYLDLJNMPFINSIQKPCSPFDVD--KPSMOKQOVKRLIQTDGKYGCT 281
 DB 1179 ISQLND---ETSTQGENSIRKKNDLEGEVAKAMKSTSEOSNLK-KSEID----- 1226
 QY 282 STQLQLEKRRKGGYC-----ECCLOKYE--DLETHLLSEQH 315
 DB 1227 ANLDQIKELKRNKNETNASLLESIKSVESEYVRIKELDECNFKREKEVSELEKLASED 1286
 QY 316 RNF-----AOSNOYOVVDIVSKLVDFVEYEKDTPKKRIRIYVGSLSPVASVYLKTEQ 371
 DB 1287 KNSKYLELOKESSEKIKELDAKTELEKIRITNLAKRKESELSR-----LKTSS 1341
 QY 372 KKKVELOHISOKDCQEDDTYVKEONFLYKETOETEKLL-----FISEPIPHSNELRGL 426
 DB 1342 EER-----KNMEBOLKIKNEIQIKNOAFERKIRKLNGSSSTIQOYSEKINTLEDE 1393
 QY 427 NEKMSNKCSTLSTADDIRONFTQPLH-----KNKQECILIDISEHTLS----- 470
 DB 1394 LIRLDONELAKETIDYRSELEKVSLSNDELLEKQNTIKSLDEILSYKDKITRNDEK 1453
 QY 471 -----ENDLEELR-----VDHYKCNIQ-----ASY 490
 DB 1454 LLSIERONKRLLESLEKQLRAQESAKAVEEGKLLEESSKEAELEKSEKMMKLEST 1513
 QY 491 HVSDFSTNDSGQPKQSDTVLFPANDLEKDLHSIFTHDSGLIT-INSSQE-----H 542
 DB 1514 IESNTELEKSSMETIRSDERLQSKSAEEDIKMLQHEKSDLLSRINESKEDIELKSK 1573
 QY 543 LTVQAK--APFHTPEEPNECDFRKNMDSLPSGKIRKRYKIILGSRKRENELEPAAEFDKRT 600
 DB 1574 LRIEAKSSGELETVKQELNNAQEKI-----RINAENTVL-RKSLIEDIERELK-DKQA 1624
 QY 601 EF-ITQENRICSSPVOST--LDLFQTSSEKSE 630

DB 1625 EIKSNOEKEKELTSRLKEFOELDSTQOKAKOSE 1658

RESULT 13
 QYAHK6 PRELIMINARY: PRF: 1173 AA.
 ID QYAHK6
 AC QYAHK6
 DT 01-JUN-2001 (TReMBrel. 17, Created)
 DT 01-JUN-2001 (TReMBrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBrel. 17, Last annotation update)
 DE LMP1.
 GN LMP1.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 NCBI_TaxID=139;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-HB19;
 RA Dunn J.J.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF305609; AKI8601.1; -;
 DR EMBL; AF305609; AKI8601.1; -;
 SQ SEQUENCE 1173 AA; 134287 MW; F0713ADF2E2311DC CRC64;

Query Match 5.2%; Score 182; DB 2; Length 1173;
 Best Local Similarity 19.4%; Pred. No. 0.0016;
 Matches 160; Conservative 144; Mismatches 290; Indels 232; Gaps 43;

QY 22 NEKNRPSLSKLTDRNPEKSKCP--LMGVF--YLDLPSTVISEKLOKIDKLG---- 72
 DB 249 NNNNTSTSLKTIKPSMSQKSELSPPSTIIGIKIRY-----SYLIKELIYELIDDIITGNV 304
 QY 73 -----GRVEEFLSKDIS-----YLINKEAKFAQ-----TLGRISPVSPFE 109
 DB 305 TLGNRLKLELIKGLSKNFKQVNELENSKNKEASNLILLIKIDLEPILNIPKDPYK 364
 QY 110 SAYVYAEFTTSPHSDG---SFKSPDVCLSRGLVKAIAKD-HDT--PNSILSNA 162
 DB 365 ETEFDLQKDKKPPHGLSKSVHSIKPIDENTKSR--QQAIKDLNEFLKNNPNDQASKT 423
 QY 163 LSWGVKILHIDIRYIEO--KKELYLLKSSSTSVRDG--KRVGSAQKTRGRKLP 218
 DB 424 LAQANKIQHLEDLSKRVHSIKPIDENTKSRQAQIKDLNEFLKNNPNDQASKTTLAQAANK 483
 QY 219 FVKVEDM-SQLY--RPPIQLIN---MPLINISIQKPCSPFDVDKPSMOKQOVKLR 270
 DB 484 IQHLEDLSKSVYSIRPIDENTKSRQAQIKDLNEFLK--NPNDQASKTTLAQAANKIOHL 541
 QY 271 IOTDGGKYGTSIQLOLEKRRKKGYCECLOKYEDELTHLL-----SEOHIRNFAOSNOYO 325
 DB 542 EDLKSQVHSIKPIDENTKSRQ-----QAIKDLNEFLKNNPNDQASKTTLAQAANKIO 593
 QY 326 VVDIVSKL-----VFDEVEYEKDTPKKR-----IK 352
 DB 594 HLEDLSKSVYSIRPIDENTKSRQAQIKDLNEFLKNNPNDQASKTTLAQAANKIOHLEDL 653
 QY 353 YSVGSLFVSASVYLKTEQ--KEVELOHISOKDCQEDDTYVKEONFLYKETOETEKLL 410
 DB 654 SKVYSIRPIDENTKSRQAQIKDLNEFLKNNPNDQASKTTLAQAANKIOHLE--DLKSVH 711
 QY 411 FISEPIPHSPNE-----LRGLNEKMSN-----KCSMLSTRAEDDIRONFTQPL 453
 DB 712 ST-RPIDENTKSRQAQIKDLNEFLKNNPNDQASKTTLAQAANKIOHLEDLSKSVYSIR 770
 QY 454 H-----KNKQECILIDISE-----HTL-----SENDELEL-----RV 479
 DB 771 DLENTKSRQAQIKDLNEFLKNNPNDQASKTTLAQAAYENNGDLKAENAYETIKTLNTQOE 830
 QY 480 DHYKCN-----QASVHVSDFSTNDSGQPKQSDTVLFPANDLEKDLHSIFTHDS 531
 DB 831 DHYKGLIRFLKRYKKEHSIESFD-----QTIKDLPKRRKALANK 869


```

OY 532 G--LITINSSOEHL-----TVOAKAFHTP-----PEEPNE-----CDFK---NMDSL 569
DB 870 GIALMLKNNKKAISFEKAIQIDKNITAYQKCIAEKNGDMQOASRNNAYLKKN 929
OY 570 PSQIHRKVKII--LGRNRENLEP---NAEPDKREF-----ITQENRICSSPVQ 616
DB 930 PNVAL--KAGIYSNMLG-NFKOSEEYLNFFNNAKMPKMEIATYNLSIKAFENNNKEESLE 986
OY 617 SLIDLFPQSEKSEFLGFTS--YTEKSGICVLDIWE---BENDSN 657
DB 987 TINKAIDNPEKSEYLYLKASINLKKNYONASISLYLVEIKNPN 1032

RESULT 14
OY 09AV25 PRELIMINARY: PRT: 1578 AA.
AC 09AV25;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PUTATIVE KINESIN-RELATED PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Buehl C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Burr P.C.,
RA Hsiao J., Zisman V., Pal G., Bowman C.L., Fujii C.Y., Vanken S.E.,
RA Bowman C.L., Craven B., Utechtack T.R., Khalak H., Feldblum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSUBA0001014 genomic sequence.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC025783; AAR20041.1; -
SQ SEQUENCE 1578 AA: 178947 MW: CBB7ELDB68A821A0 CRC64;

Query Match 5.2%; Score 181.5; DB 10; Length 1578;
Best Local Similarity 19.9%; Pred. No. 0.0025;
Matches 152; Conservative 109; Mismatches 256; Indels 245; Gaps 38;

OY 19 QVKNKNNKPLSKLK---TDNRPEKSKCKPLMGKVFYLDLPVTTISEK--LQDKDKLGG 73
DB 340 QKRTKRLKQGLTETLKECADSKQSELE-----LEL-SISISESDSLKQTEIEELKR 390
OY 74 RVEEFLSKDILSYLISNKKKFAOGLGRISPVSPESAYTAETTSRPHSHDSSFKSPDT 133
DB 381 SLEEVYTA-----QTISSR-----SPRSGDAITELQK--EVEDDVQFLKESN 428
OY 134 VCLG-----RGLLVEKA-----IKDHDFIPSNLSLNL 163
DB 429 ASLATOAKAOEANIETVSIQLEETIEVQRAEISNLSHTSDLDHEVSPNNLLIQEDV 488
OY 164 SWGVAT-LHIDDIRIYIOKKKELLYLKKSSIVDGGKRVSGQAKRTGKRLKPPYKV 222
DB 489 EMARVYSLKEDEILMLRKRIKIDRMHLV-----ENPNGRSSGA----- 524
OY 223 EDMGSLYPPFYQL-TNMPFINYSIO---KPCSPDVADKPPSMQOTQY----- 267
DB 525 -----TYLELEKEDFLKVKIQLELEKDCSELTDLELITYLKEKSEVAKEDPSV 575
OY 268 --KLTIOGTGDKYGGTSLQQLKERRKKGCCLOKYEDELTHLL-----SEQHRNPAQ 320
DB 576 PNSEVSSEGD-----LSDRLTSKYK--YLE--TKCADLELKLISPRSESELEEKQ 624
OY 321 SNQYQVVD--DIVSKL--VFDFVEYEKDTPKPKRIKYSVGSLSPVASVYLKTE 370
DB 625 KSOEKLKRILELSDLRKLKLSGFHALMEEGDTDSAKSKLSE-----KIDE 672
OY 371 QKEKVELHISQKDDQEDDTYVKEQNFLYKETOETEKTL-LFISDFPHSPNELLGLINEX 429

```

```

DB 673 NDNKTEL-----DALRSTVLLKBEI--ESLQSKKEMESPISB-IMNKKNKLEELLE 723
OY 430 MSNCKSMUSTAEEDDIRQNFQPLKKNQOECLDISEHTLSBNDLEELRVHYKCNIOAS 489
DB 724 SLKECSITACACIDEMREBELLL-----TSSIDSHVSTNNVLETITELESQVMLE 774
OY 480 VHVSDPSTDN-----SGSQPKQKSDVLPFPAKDLKEDKLSITTHSGITINSSQEH 542
DB 775 LHSKLEHENVLSEFTIGLSOL--TYL-----ANKELSMQMDERSRLTNLKDDEL 826
OY 543 LTVQA-KAPFTPEEPNE--CDFKN-MDSLPSGKIHKKVITIGR---NRKENLFE--- 591
DB 827 EGVENQKVELKQMDERSRLITNLKDELEQYEAQKVELKQMDERSRLITNLKDELEVE 886
OY 592 -----PNAEPDKREFITQENRICSSPVOSLDLF 622
DB 887 AQVELKENOLESRRRLSEVOEDSEALRRSNAKLOATVDHVEE---CKSIQITVADL- 941
OY 623 QTSSEKSEFLGFTSYTEKS-----GICNVLDIMEENS 655
DB 942 --KKQLEVHGVAHLEQLEQSKRKRTMDFCFTLESLAKUS 981

RESULT 15
OY 049545 PRELIMINARY: PRT: 1344 AA.
AC 049545;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMblrel. 08, Last annotation update)
DE LMP1 AND LMP2 GENES.
OS Mycoplasma hominis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmales; Mycoplasma.
OX NCBI_TaxID=2098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG21;
RX MEDLINE=95104998; PubMed=7806360;
RA Ladefoged S.A., Birkelund S., Hauge S., Brock B., Jensen L.T.,
RA Christiansen G.;
RT "A 135-kilodalton surface antigen of Mycoplasma hominis PG21 contains
RT multiple directly repeated sequences.";
RL Infect. Immun. 63:212-223(1995).
DR EMBL: X81475; CAAS7228.1; -
SQ SEQUENCE 1344 AA: 149404 MW: 6E0105F7365AF8CF CRC64;

Query Match 5.0%; Score 177; DB 2; Length 1344;
Best Local Similarity 19.1%; Pred. No. 0.004;
Matches 132; Conservative 132; Mismatches 256; Indels 170; Gaps 30;

OY 19 QVKN-----EKNRPSLSKTDNRPESKCKPLMGKVFYLDLPVTTISEKLO----- 65
DB 221 ETKMANQAVASNNITASQMSANSS-----LDKAAETITKLETTNNKKE 273
OY 66 --KIDKLGKVEEFLSKDILSYLISNKKKFAOGLGRISPVSPESAYTAETTSRPHS 122
DB 274 AKFNELKQTRNIOGFIN-----TNKNNPYSLSIQLTGSKRSKSNVY-----DS 319
OY 123 HDGSEFKSPDT--VCLSRGL-----LVKKAIDHDFIPSNLSLNLNLSMGVATLHID 173
DB 320 SNKSDIESANTELKQALAKANADKVOADNLAKSIKEO-----LNNSVSNANLTSARLTQKD 375
OY 174 DIRRYIEQKKKELYL-LKSSSVVBDGGKRVSGQAKRTGRLKPPYKVEDMGLYRPF 232
DB 376 NT--IQQKTELEKEVQKADQATKSNMTASQMSAKSSIDAKVAF----- 417
OY 233 YLQLTNMPFINYSIQKPCSPDVADKPPSMQOTQYKRLQ-----TDGKYGTSIQLO 286
DB 418 -----ITKLTETPKDKKAEKFNELKQTRNIOGFINTNNKNNPNVSELISQLT 464

```

Fri Dec 28 08:22:43 2001

us-09-830-647-1.rpt

Page 10

```
OY 287 LKRRKKGYCECLOKVEDLETHLLSEOHNFASNOYOVDIVSKLVDPFVEYEKDP 346
DB 465 SKRDSKNSVTD--SNKSDIES-ANTELKQALAKANADKVOADNLAKSI-----510
OY 347 KKRRIKYSVGLSPVSAVYLK--TEQKEVELQHSQKDCQEDPTVKEDNFLYKETOE 404
DB 511 -KEOLNNSVSANANTLSAKLTDKDNTOQAKTEL---EKEVQADQAIKSN--TAS 560
OY 405 TEKRLFISEPIHPSENLRGINERKMSKCSMLSTAEEDIRQNTQOLPLAKKQECJDI 464
DB 561 MOSAKSWLDKVAETIKKLETFENKDKREKFNELKQTRNOIQEFINTKNKNNPYSELISOL 620
OY 465 SEHTLSENDELELRVDHYKCNIOASVHVSDFSTDNS-----GSQPRQKSDTVLLEPAK 516
DB 621 TSKRDSKNSVTD-----SSNKSDIESANTELKQALAKANADKVOADNL--AK 665
OY 517 DUKEDLHSITFDGLITINSSQEHLTVOAKAPFHPPEEPNEC-DFKNMDSLPSGK--573
DB 666 SIKED-LNNSVSANANTLSAKLTDKDNTOQAKTELEKEVQANQAIKSNNTASWOSAKSS 724
OY 574 IHRKYKIIIGNRKENLEPNAEFD--KRT---EFTTOENRIGSSPVQOSLLDLQOTSE 626
DB 725 LDQKVAETIKKLETFENKDKREKFNELKQTRNOIQEFINTKNK--NPVYSELISOLTSK 780
OY 627 EKSEFLGFTYTEKSGICNVLDIWEENSND 656
DB 781 RDSK---NSVTDSS--NKSDI-ESANTE 802
```

Search completed: December 27, 2001, 16:55:40
Job time: 240 sec